QUESTIONS from questionnaire to use

2015 to 2019 are comparable

85 = bad data – remove

Age only over 17 – remove younger

(Code - Subject – response numbers)

Demographics

AGE2 - Age – 7 to 17

~~NOMARR2 - Number of times married – 1, 2, 99~~ (3/5 have skipped this question)

IRMARIT – marital status – 1 to 4 - added

SERVICE - Ever been in the US armed forces– 1, 2, 99

HEALTH - Overall health – 1 to 5, 94

MOVSINPYR2 - Times moved in the past 12 months – 0 to 3

SEXATRACT - Sexual attraction – 1 to 6, 94+97+98+99 – 1 to 6, 98

SEXIDENT - Sexual identity – 1 to 3, 94+97+98+99– 1 to 3, 94, 98

DIFFTHINK - Serious difficulty concentrating, remembering, decision – 1, 2

DIFFWALK - Serious difficulty walking – 1, 2

DIFFDRESS - Serious difficulty dressing or bathing – 1, 2

DIFFERAND - Difficulty doing errands alone – 1, 2

IREDUHIGHST2 – Highest Education – 1 to 11 - added

Work

WRKSTATWK2 - Work situation in the past week – 1 to 9

~~WRK35WKUS - Work 35 or more hours per week – 1, 2, 99~~ (almost half skipped the Q)

~~WRKDPSTYR - Past 12 months worked any job – 1, 2, 97+99~~ (most skipped this Q)

WRKNUMJOB2 - Past 12 months how many employers – 1 to 4

IRWRKSTAT – employment status – 1 to 4

Income

IRPINC3 - Income – 1 to 7

CIG100LF – 1+3+5, 2, 91 – Have you smoked 100cigs in your life? – from partdataset

CIGTRY – range 1 – 65 – Age when first smoke cigarette

ALCUS30D – Estimate number of drinks per day that you drink: # (range (1 – 85) OR n/a (991, 993)

OUTCOME

(Code – response – question)

*Cigarette use*

Learning from:

CIGEVER – 1, 2 – Ever smoked a cigarette? # If no, grey out smoke prediction

~~CIGWILYR – 1 to 4, 94+99 – Do you think you will smoke cig in next 12 months?~~

CIGDLYMO – 1+5, 2, 91 – Ever smoked cig every day for 30 days?

Outcome:

CG30EST - Estimate days to smoke cigarettes in the next 30 days: .. days (1 to 6) OR n/a (91, 93)

*Alcohol use*

Learning from:

ALCEVER – 1, 2 – Ever had a drink of alcohol beverage?

ALCREC – 1+11, 2, 3, 8, 91 – Time since last drank alcoholic beverage

ALCYRTOT – range 1- 365, 991, 993 – Total # days used alcohol in past 12 mos

Outcome:

ALCDAYS - Estimate days to drink alcohol beverages in the next 30 days: .. days (range 1 – 30) OR n/a (91, 93)

How likely have you used or are you going to use the following drugs:

MJEVER – 1, 2 - Marijuana/ Hashish - %

COCEVER – 1, 2 – Cocaine - %

CRKEVER – 1, 2+91 – Crack - %

HEREVER – 1, 2 – Heroin - %

LSD, PCP, PEYOTE, MESC, PSILCY, ECSTMOLLY, KETMINESK, DMTAMTFXY, SALVIADIV, HALLUCOTH – 1+3, 2+91 – Hallucinogens - %

AMYLNIT, CLEFLU, GAS, GLUE, ETHER, SOLVENT, LGAS, NITOXID, FELTMARKR, SPPAINT, AIRDUSTER, OTHAEROS, INHALOTH – 1+3, 2, 91 – Inhalants - %

METHAMEVR – 1, 2 -Methamphetamines - %

PNRANYLIF – 1,+5, 2 - Pain relievers - %

TRQANYLIF – 1+5, 2 Tranquilizers - %

STMANYLIF – 1+5, 2 – Stimulants - %

SEDANYLIF – 1+5, 2 – Sedatives - %

*Mental health – K6*

During the past 30 days how often did you feel:

DSTNRV30 – score 1 to 5 – nervous

DSTHOP30 – score 1 to 5 – hopeless

DSTRST30 – score 1 to 5 – restless/ fidgety

DSTCHR30 – score 1 to 5 – so sad or depressed that nothing could cheer you up

DSTEFF30 – score 1 to 5 – that everything was an effort

DSTNGD30 – score 1 to 5 – worthless

Score 6-18 No probable serious mental illness

Score 19-30 Probable serious mental illness

?? Go here for an actual test -

2019 (data1\_df\_cp)

Total rows/ columns: 56,136/ 2,741

Columns used: 64

Rows after age removal: 42,739

Rows after error 85 removal: 41,950

Rows after removal of some responses:

Final dataframe rows/ columns: 24,584/ 64

2018 (data2\_df\_cp)

Total rows/ columns: 56,313/ 2,691

Columns used: 64

Rows after age removal: 43,026

Rows after error 85 removal: 42,252

Rows after removal of some responses:

Final dataframe rows/ columns: 24,575/ 64

2017 (data3\_df\_cp)

Total rows/ columns: 56,276/ 2,668

Columns used: 64

Rows after age removal: 42,554

Rows after error 85 removal: 41,761

Rows after removal of some responses:

Final dataframe rows/ columns: 23,961/ 64

2016 (data4\_df\_cp)

Total rows/ columns: 56,897/ 2,668

Columns used: 64

Rows after age removal: 42,625

Rows after error 85 removal: 41,844

Rows after removal of some responses:

Final dataframe rows/ columns: 23,613/ 64

2015 (data5\_df\_cp)

Total rows/ columns: 57,146/ 2,679

Columns used: 64

Rows after age removal: 43,561

Rows after error 85 removal: 42,752

Rows after removal of some responses:

Final dataframe rows/ columns: 23,810/ 64

(data\_complete\_df)

Final total dataframe after merging:

Rows/ columns: 120,543/ 64